

Change in Digital Business: Chatbots and Their Significance on Online Purchasing

Author: Azad-Asnafor Gelici
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,

The world of business is constantly changing, the trend is currently on online business. However, whatever the trend is, support is always necessary. The kind of support which fits best with the current business format, is support through chatbots. As the number of e-business customers needing support has reached a level that human support for every customer is not possible anymore. As the usage of chatbots in online business is fairly new, improvements are still necessary. One of those improvements is the ability to create more humanlike chatbots. The variables Competence and Empathy, relating to chatbot humanness, were compared to the variable Likability, this in turn was compared to Consideration. The analysis showed that these comparisons showed for positive causal relationships. This enables future research regarding chatbots for further development of its human traits and supports theories regarding anthropomorphism chatbots.

Graduation Committee members: Dr. C. Herrando & Dr. E. Constantinides

Keywords

Chatbots, Customer perception, Customer support, Anthropomorphism, Artificial Intelligence, Customer Satisfaction, Robotic Development

1. INTRODUCTION

As time progresses the importance of digital business becomes more apparent, many companies and organisations switch from their physical stores and services to a combination or to the sole use of online business (Changchit 2006). As every customer/consumer needs some form of support, also digital business needs to set up a well-functioning digital support to aid these consumers. At first this was done through online chats with real people, then a switch towards artificial intelligence(AI) chatbots had arisen as the number of digital business customers in need of support has reached a level that does not see human support for every customer as a possible scenario anymore. So instead a switch needed to be made to a more reliable method of supporting customers within a certain time frame that keeps them satisfied with the support, this is through chatbots (Thomas, 2016). According to Zumstein and Hundertmark (2017) the number of chatbots strongly increased over the last few years. Nevertheless, 84% of internet users still have not used a chatbot (Jain et al. 2017). This paper also states that chatbots that displayed human-like conversation were preferred. As Adiwardana et al. (2020) state "...users expect chatbot systems to behave and communicate like humans. If the chatbot is seen to be "acting like a machine", it is deemed to be below standard. It is required to have the same tone, sensitivity and behaviour than a human...". Thus a problem that arises from this switch to chatbot support, is the correct development of chatbots to be able to come across more human. Even though, chatbots have developed and are better able to understand the customers, however still, room for improvement is still apparent. As research is still being performed focusing on creating more human-like chatbots (Adiwardana et al., 2020).

The academic relevance this research brings is the importance of human emotions related to the interaction with Artificial Intelligence, gathering information that gives further understanding of human interaction with AI will create more relevant literature corresponding to human and AI relations. Furthermore, knowing which human emotions arise when humans interact with AI and strengthening the aspects of the AI chatbot related to the rise of positive human emotions, will lead to the better performance of chatbots and thus a better performance of companies and organisations using these chatbots. Moreover, weakening the aspects of the AI chatbot related to the rise of negative human emotions will be beneficial towards the same goal.

1.1 Research Question

This research tries to find out how customers perceive AI chatbots and also how they relate these bots towards human chat. Eventually after knowing the customers' dissatisfaction different options can be recommended to improve the quality of an AI chatbot. The focus is not the technical aspects, such as coding, that make up the chatbot, but primarily the focus on human conversational features, that a chatbot might lack or is weak in and needs improvement for.

The research question can be used as an indicator of achieving the research objective, whenever a sufficient answer to the research question is created, the objective is achieved. For this study the research question is: Looking at human traits, how can Artificial Intelligence chatbots be adapted to improve customer satisfaction with said chatbot?

2. THEORETICAL FRAMEWORK

Relating to the literature discussed in the "literature review" section, sensibleness, specificity, affect and emotion can be seen as aspects of a human-like chatbot (Banchs, 2017; Adiwardana et al., 2020).

As in this research, competence and empathy are both created from the fusion of their sub variables: sensibleness, specificity, affect and emotion. Competence and empathy can be used to identify human-like chatbots. Thus the connection of "Competence" and "Empathy" with "Human-like Chatbot" is explained in the research model.

Now that the makings of the "Human-like Chatbot" variable is explained we can theorize its influences, This research tries to show that the more human-like a chatbot is the higher its likability is; this is supported by research (Bartneck et al, 2009; Epley et al., 2007; Złotowski et al., 2015; Moussawi et al., 2020; Go & Shyam Sundar 2019; Kiesler & Goetz 2002). In the model this is indicated with the relationship between "Human-like Chatbot" and "Chatbot Likability".

This research tries to determine whether or not a more human chatbot improves its likability, however, what if it does? Is that the end of the chain of influence? No, it is not. In this research there is a final variable that will show the success a more liked chatbot brings with it; this is the variable "Chatbot Consideration". This variable indicates how much the actions suggested or advice given by chatbots is actually considered by chatbot users. In the research the relation between a chatbot's likability and its consideration is measured. The more likable a chatbot, the more it is considered (Cialdini, 1993; Kenrick, Neuberg, & Cialdini, 2002; Roskos-Ewoldsen & Fazio, 1992).

The understanding of the variables and the references mentioned will be further elaborated upon in the next section "Literature Review".

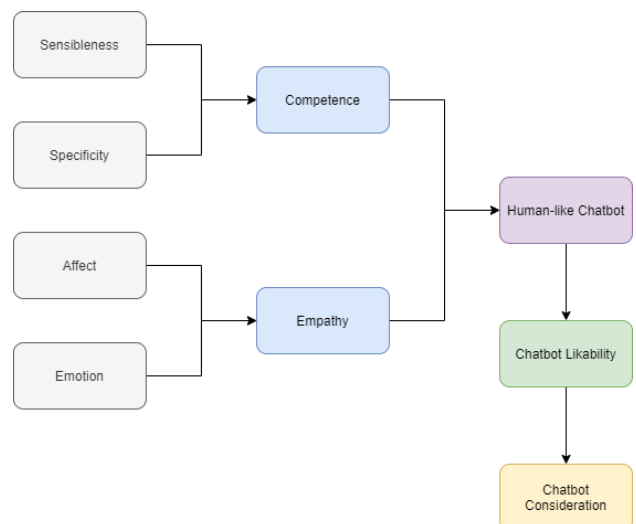


Figure 1. Research model

3. LITERATURE REVIEW

This literature review enables a better understanding as to what is needed to improve customer satisfaction with chatbots. This review elaborates as to why chatbots should be humanized to increase customer satisfaction, furthermore, the variables relating to human-like chatbots will be mentioned and elaborated upon.

3.1 Anthropomorphism

It is a human tendency to attribute human traits to non-human entities, this is what is called anthropomorphism (Epley et al., 2007; Zlotowski et al., 2015). Further Epley et al. (2007) describe anthropomorphism as the tendency to connect humanlike characteristics, motivations, emotions, or intentions to the behavior of nonhuman agents. As chatbots are obviously not human, according to this theory it would be preferred by humans for the chatbot to have human traits. Furthermore anthropomorphism positively impacts the intention to adopt personal intelligent agents (Moussawi et al., 2020). Go & Shyam Sundar (2019) also state that an intelligent and humanized machine is able to have a better engagement when communicating with a human. Looking at the proposed research model this would explain the positive connection between “Human-Like Chatbot” and “Chatbot Likability”. The findings of Kiesler & Goetz (2002) also support this correlation, they found that people were more likely to cooperate with playful, “human”, robots than with serious, “robotic”, robots. Bartneck et al. (2009) also state that a highly anthropomorphic and intelligent robot is more likely to be perceived as more likeable.

Now that this connection has been established a way needs to be found to measure how human-like a chatbot really is, variables are needed to measure this. In the following subsection these variables will be covered in detail.

3.2 The variables

In order to judge the perceived level of human traits a chatbot possesses, the traits need to be specified. Inherent properties of human-human communication and interaction are affect and emotion (Banchs, 2017). These can thus be used as traits a human-like chatbot should possess. Additionally Adiwardana et al. (2020) suggest that sensibleness and specificity are also two rudimentary aspects of a human-like chatbot. These traits can be used in the research as variables which measure the “humanness” of a chatbot, these variables will be elaborated upon further. Apart from variables that measure this humanness, variables are also necessary that explain the likability of a chatbot, this in order to show the correlation between chatbot humanness and chatbot likability. For this paper these variables will be Likability and Consideration. The likability variable refers to the positive impression of a person or a thing (Bartneck et al. 2009), in this research thus the positive impression of a chatbot. The consideration variable will be created to confirm the true likability of chatbots, as a message will be considered more strongly when there is a higher likability of the source of said message (Roskos-Ewoldsen & Fazio, 1992).

Now the variables are known, however, what do the variables entail, what do affect, emotion, sensibleness and specificity mean? What do likability and consideration mean?

Affect

Affect refers to the fundamental affective experience of feeling; it is a mental state, it is about how we feel (Hogg, Abrams & Martin, 2010). Creating items for the survey we need items that make clear how the person feels regarding a chatbot.

Emotion

According to Shouse (2005) emotion is regarded as the direct expression of affect. Thus for the survey we need items that

show the noticeability of emotion expressed by the chatbot and the ability the chatbot notices your emotional expression.

Empathy

Affect and emotion can both be related to a single term, empathy. In empathy, people feel the emotions of others, this grants it by definition both aspects of affect and emotion (Pijnenborg et al. 2012).

To simplify, the model we can thus, add both affect and emotion in the same variable, empathy.

Sensibleness

Relating to sensibleness, it can be seen as following, according to Adiwardana et al. (2020), sensibleness would mean, being sensible, being able to converse with a human being while making sense in the context. Items should be created focused around this definition. They also suggest specificity to be getting an answer that is specific to the question and not a widespread answer applicable to many questions. Items should be created that address whether a bot is specific or not.

Specificity

Relating to specificity, it can be seen as following, according to Adiwardana et al. (2020), specificity can be defined as to be getting an answer that is specific to the question and not a widespread answer applicable to many questions. Items should be created that address whether a bot is specific or not.

Competence

As the variables sensibleness and specificity relate to the proper working of a chatbot, again for simplification, they can be connected to the competence variable. Competence is in simple terms the ability to perform well (Weinert, 2001).

Likability

It is in human nature whenever there is similarity to ourselves, likability is increased (Byrne, 1971; Carli, Ganley, & Pierce-Otay, 1991; Hogg, Cooper-Shaw, & Holzworth, 1993). Looking at the research this paper approaches, as humanlike chatbots are more similar to humans than non-humanlike chatbots, humanlike chatbots should be more likable.

Consideration

Consideration refers to the consideration of following the advice given by chatbots. Cialdini, (1993); Kenrick, Neuberg, & Cialdini, (2002), label liability as a persuasion factor. Thus this links consideration and likability together, when a chatbot is likable their messages will be considered more. This can be supported by experiments taken by Roskos-Ewoldsen and Fazio (1992), they found support for this hypothesis. According to them, a message will be considered more strongly when there is a higher likability of the source of said message.

3.3 Hypotheses

The variables mentioned in this literature review are fundamental to the research, these variables were used to measure the soundness of the following hypotheses.

H1. A more human chatbot increases the chatbot’s likability.

Competence and empathy, created from their sub-variables sensibleness, specificity, affect and emotion, are human traits that can be associated with chatbots and make them seem more human (Adiwardana et al. 2020; Banchs, 2017). This human

chatbot can be used to determine the hypothesis: “A more human chatbot increases the chatbot’s likability” (Bartneck et al., 2009). Support for this hypotheses would mean the existence of a positive causal relationship between the variables competence and empathy with likability. Also anthropomorphism establishes that giving chatbots human traits increases its likability (Epley et al., 2007; Złotowski et al., 2015).

H2. A more likable chatbot will be considered more strongly.

Furthermore, another hypothesis can be thought of. This is, whenever a chatbot is more likable its advice will be more considered. The hypotheses is: “A more likable chatbot will be considered more strongly”. This is important for the world of digital business as the chatbots need to fulfill their function of supporting the customers. This hypotheses is supported by the theories of Cialdini, (1993); Kenrick, Neuberg, & Cialdini, (2002); Roskos-Ewoldsen and Fazio (1992). This research supports this hypotheses when the data shows that there is a positive causal relationship between chatbot likability and chatbot consideration.

4. METHODOLOGY

The research question this paper aims to answer is: Looking at human traits, how can Artificial Intelligence chatbots be adapted to improve customer satisfaction with said chatbot? To be able to answer this question, the opinions of chatbot users have to be known, the opinion relating to how human chatbot users perceive chatbots and the opinion relating to how much users like chatbots. The research is conducted through a survey. Quantitative data allows for the understanding of the opinions of the applicants and how this could relate to an improvement of chatbots.

4.1. Research Design

The survey is designed through the usage of statements, a sentence which exists of an opinion related to a subject, statements were created solely for this research that represent the essence of the variables mentioned in the literature review section. The survey will measure 6 variables, each of which gets measured through 3 questions. The variables are, sensibleness, specificity, affect, emotion, likability and consideration. In turn, the variables sensibleness and specificity will be fused into competence and affect and emotion will be fused into empathy. The statements will be formulated through a 7 point likert scale, this means that the statements have 7 possible answers with their own given value, with 1 being “strongly disagree” and 7 being “strongly agree”.

When an item is formulated negatively, where necessary the values will be flipped in order to receive data that over all different variables provides values that have the same meaning. In practice this would mean, rating a statements with a 7 and thus strongly agreeing, could with a negative formulation mean, strongly disagreeing with the proposed variable, flipping the 7 value to a 1 value would then correctly represent the opinion of the applicant. This is called reversed coding. In this research this would make the analysis simple, the higher the mean score the higher the rating is of a variable by people. For the analysis of the survey two statements’ values will be reversed to be able to receive correct data, these are related to the variable specificity, the statements are: “I often get generic answers such as yes and ok.” and “I feel like my questions are not answered satisfactory.”. As answering Strongly agree to these

statements would have the opposite meaning towards the perception of chatbots compared to all the other statements.

From the survey a better understanding of how applicants feel about the chatbot’s competence, subdivided in sensibleness and specificity, empathy, subdivided in affect and emotion, its likability and the consideration will be gained. From this data a determination can be made regarding the level of perceived sensibleness, specificity, affect and emotion of the applicant, also the likability will be measured, with these measurements a correlation is measured, the correlation of competence (sensibleness and specificity) and empathy (affect and emotion) with likability. The consideration variable will be used to measure whether or not there is a correlation between the chatbots’ likability and the consideration of following its advice.

4.2 The Survey

The survey was created with the aim of measuring chatbot likability and perceived chatbot humanness. Furthermore, the survey consists of questions regarding personal data, gender, age, education level, employment situation and nationality. Apart from this, this survey has also the aim of measuring how many of the applicants know what chatbots are or have any knowledge of chatbots. Towards the aim of receiving reliable data, the question “Do you have any experience with or knowledge of chatbots?” will be asked before the 18 statements are shown. Any applicant which answers yes will continue with the rest of the survey. Any applicant which answers no, will be withheld from the rest of the survey and is finished with their application. This question is crucial towards the success of this research, as it allows for more reliable data and provides important data related to this research topic. The survey design can be found in Appendix 12.1.

5. RESULTS

This section focuses on the results of the survey and the analysis of it, in the discussion section these results will be further discussed and elaborated upon.

5.1. Sample Description

The survey reached a total of 151 applicants. This survey has 5 applicants which did not answer a single question, this brings the actual amount of useful applicants to a total of 146. Of the applicants, 80,1% claim they have experience with or knowledge of chatbots, thus 19,9% state they have no knowledge of chatbots. This leaves 117 people which have continued with the rest of the survey. Furthermore, 46,2% of applicants, 54 people, strongly disagree with the statement “I prefer chatbots over human chat.”, the mean score regarding this statement is 2,30, laying closely to disagree. This is also the lowest mean score measured in this research. The highest mean score belongs to the statement, “I believe the information given by chatbots to be true.”, with a score of 4,79. All the other results can be found within Appendix 12.2.

5.2. Analysis Results

Combining the variables together to the pre-mentioned variables “competence”, “empathy”, “likability” and “consideration”, the mean scores, standard deviations and Cronbach alpha’s look as followed:

Table 1. Variable Statistics

	Competence	Empathy	Likability	Consideration
Mean	3,8262	2,9060	3,3362	4,0513
Standard Deviation	1,00773	1,19172	1,33225	1,26219
Cronbach's Alpha	0,784	0,832	0,764	0,718

As stated by Pallant (2010), a Cronbach's Alpha higher than or equal to 0,7 states that the variable is reliable, furthermore when the variable has less than 10 items, the Cronbach's alpha should be more than 0,5, as the variables are higher than 0,5 and even higher than 0,7 these variables are reliable and internally consistent.

A valuable source of information apart from the main research goal is related to the item "Do you have any experience with or knowledge of chatbots?". This item tells us that 80,1% of the applicants have knowledge of or have experience with chatbots.

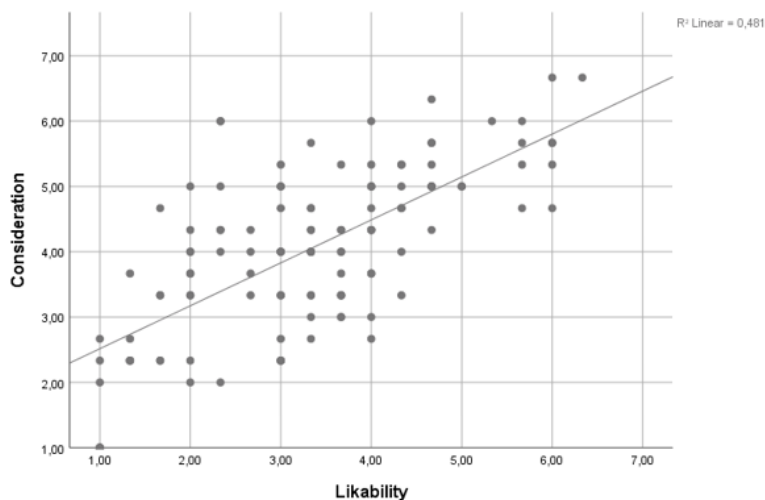
Looking at the correlations which can be set up through this survey, there are some which will help towards answering the research question. In this research the strength and direction of a relationship between two variables are measured, thus we will use the Pearson correlation. In Table 2. the variable's Pearson correlations are noted.

Table 2. Variable Pearson Correlations (N = 117)

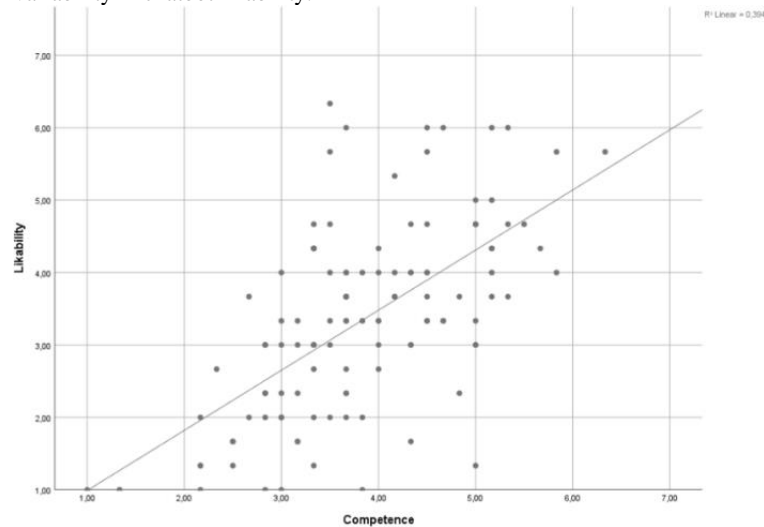
	Empathy	Likability	Consideration
Competence	0,33**	0,63**	0,59**
Empathy		0,59**	0,47**
Likability			0,69**

Note. **Correlation is statistically significant at the 0,01 level.

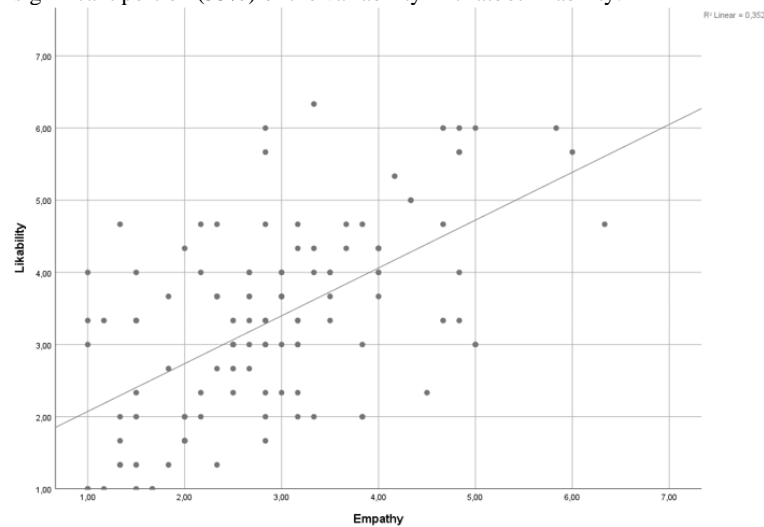
To support the hypotheses that higher chatbot likability leads to higher consideration of following the recommendations provided by chatbots, a Pearson correlation was conducted. Likability was strongly and positively related to consideration. $R(115) = 0,69$, $p < 0,001$. The correlation is shown in Table 2 and the scatterplot can be found in Figure 2. The effect size for likability ($r^2 = 0,48$) indicates that the level of likability a user experiences related to chatbots accounts for a great portion (48%) of the variability in chatbot consideration.

**Figure 2. Likability and Consideration Scatterplot**

These values indicate that there is a strong correlation between the variables likability and consideration. As now, it is known that chatbot likability indeed leads to more consideration of following the advice given by chatbots. A different correlation needs to be set up to confirm that more human chatbots lead to improvement of the chatbot likability. Chatbot likability does not only arise from liking the usage of chatbots, but also from having an useful experience with a chatbot. Again, a Pearson correlation was conducted, now, to examine the relationships between chatbot competence and empathy related to the chatbot likability. Chatbot likability was more strongly positively related to competence, $r(115) = 0,63$, $p < 0,001$, compared to empathy, $r(115) = 0,59$, $p < 0,001$. These correlations can be found in Table 2 and the scatterplots can be found in Figure 3 and Figure 4. The effect size for competence ($r^2 = 0,39$) indicates that the level of competence that the customer believes the chatbot to have accounts for a considerable portion of the variability in chatbot likability.

**Figure 3. Likability and Competence Scatterplot**

This portion is slightly bigger than the one relating to empathy. The effect size for empathy ($r^2 = 0,35$) indicates that the level of perceived chatbot empathy by the customer accounts for a significant portion (35%) of the variability in chatbot likability.

**Figure 4. Likability and Empathy Scatterplot**

Even though these correlations are somewhat weaker compared to the correlation between likability and consideration, this data is significant enough to state that competence and empathy are human factors that if present will improve the likability of

chatbots, this in turn raises the consideration factor experienced by customers regarding the advice given by chatbots.

The Pearson correlations between “competence” and “likability”, “empathy” and “likability” and “likability” and “consideration” are all positively correlated and statistically significant.

From the data it can be taken that young adults, aged 21 -25, are least of the opinion that chatbots are disliked, this is related to the variable likability, compared to the other age groups, see Appendix 12.4 Case Summaries for this and the following data, with a mean score of 3,4749. The oldest group the survey reached, aged 36 – 50, are mostly of the opinion that chatbots are underperforming compared to the other groups, with a mean score of 2,4286. There is are no significant differences between genders and education levels regarding the variable.

6. DISCUSSION

In this section the results of the research and the implications will be discussed, the results mentioned will be crucial towards answering the research question, this will happen in the next section, “conclusion”.

As stated in the results section 80,1% of the applicants have knowledge of or have experience with chatbots. Thus, chatbots are not something reserved for niche markets as this percentage indicates that implementing chatbots in your company does not come as something alien towards your customers, a majority of people know already what chatbots are. Looking back at research provided by Jain et al. (2017), in 2017, 84% of internet users still had not used a chatbot. This is a significant change, according to this research currently in 2020 only 19,9% of people have no knowledge of chatbots. This supports the research done by Zumstein and Hundertmark (2017) who claim that the number of chatbots strongly increased over the last few years.

The results section shows us the correlations that can be taken from the data analysis, this allows us to determine whether or not there is truth towards the hypotheses of this research. The hypothesis: “A more human chatbot increases the chatbot’s likability” (Bartneck et al, 2009). Is supported by the results, as the results indicate there is a positive causal relationship between the variables competence and empathy with likability.

Furthermore, the hypothesis: “A more likable chatbot will be considered more strongly”. Is also supported by the results of this research. This is because there is a positive causal relationship between chatbot likability and chatbot consideration.

Now that is known that the set up hypotheses are correct the research question can be answered. For this a further discussion is necessary regarding whether an improvement is even necessary. The following part of the discussion will focus on this.

The results section has shown the correlations that can be taken from the data analysis, now however, the actual values need to be analysed to figure out whether or not an improvement regarding chatbot humanness and likability is necessary. During the research, items were rated from strongly disagree, to strongly agree, corresponding with the values 1 through 7. These values have been set up that a score lower than 4 indicates a perceived underperformance regarding the variable

and a score higher than 4 a perceived well performance regarding the variable. The more skewed towards the extremes 1 and 7 the stronger the opinion. Thus looking at the mean score regarding the variables shows the average value/opinion people have of the variable. In appendix 12.2 Survey Data, under “mean scores per variable”, these values can be found. Looking at Table 1 you can see the combined variables’ mean scores. The variables competence and consideration score rounded a 4, this would mean that the applicants neither perceive chatbots, relating to these variables, to be underperforming or performing well, they are neutral. This means that there is room left for improvement regarding these variables. Furthermore, empathy and likability score about a 3, this means that the applicants perceive chatbots to be somewhat underperforming relating to these variables. Thus there is definitely room for improvement regarding these variables. As there only is a moderate correlation between competence and likability and empathy and likability. This suggests that there are other variables apart from variables related to humanness that explain this underperformance of chatbots compared to human chat.

7. CONCLUSION

This research focuses on the research question: “Looking at human traits, how can Artificial Intelligence chatbots be adapted to improve customer satisfaction with said chatbot?” To answer this question, a research was set up with the design focused on a survey. The survey delivered valuable results towards answering the research question. The data allowed for a confirmation that competence and empathy are positively correlated to likability and that likability positively correlates to consideration. The data also shows that the applicants do not perceive chatbots to be performing well regarding their competency and empathy. They also do not like chatbots and do not consider following their recommendations. Overall regarding all the variables, there is a small dissatisfaction related to it. Increasing the chatbot humanness leads to an improvement in chatbot likability, this in turn increases the consideration of following advice given by chatbots. However, improving the humanness is not the only way to improve chatbot likability. As the moderate correlation does not fully account for the strong improvement necessary to make customers satisfied with chatbots, to the degree that they prefer it over human chat. As empathy was rated the lowest from the variables, this should first be improved within chatbots to improve the chatbot likability, further, competence also has room for improvement to improve the likability. The next level for robots is ahead, it only needs to become more human.

8. IMPLICATIONS

This research provides a new insight for the development of chatbots. From the viewpoint of humanlike chatbots, this research shows that a more human chatbot enables a better likability of the chatbot and in turn a higher consideration of following advice given by the chatbot. Companies can learn from this research and start the development process of making their chatbots even more human. These improvements need to be made regarding the chatbot’s competence and level of empathy. Businesses need to implement chatbots or update their current chatbots to the state that they are more human,. This will lead to a higher customer satisfaction as they will like the chatbots more and will follow the advice given more frequently. Further, this research supports the past research done on anthropomorphism and reinforces the theory that applying human characteristics to a nonhuman entity makes it more likable (Epley et al., 2007; Złotowski et al., 2015; Moussawi et al., 2020; Go & Shyam Sundar, 2019; Kiesler & Goetz, 2002).

This research also supports the theory of Roskos-Ewoldson & Fazio (1992), they suggest that a message will be considered more strongly when there is a higher likability of the source of said message. The relationship between likability and consideration in this research shows that there is truth to this theory.

9. LIMITATIONS AND FUTURE RESEARCH

This section focuses around the limitations this research experienced which did not allow for the full realisation of the originally planned method of research also recommendations are done to address future lines of research to further develop upon chatbots.

9.1 Limitations

A big limitation this research came across was at the very start, this research was conducted during the unfortunate outbreak of the COVID-19 pandemic, with this outbreak many limitations were issued to the people around the world. In the Netherlands the universities were closed resulting as a measure to prevent the spread of the disease (Algemene Zaken, 2020). This prevented the use of neuroscientific tools available at the University of Twente to be used for the original method of research, instead of analysis through a survey, analysing neurologic responses, using the mentioned tools, to chatbots would have been analysed to determine the applicants' perception of chatbots. Another limitation was the lack of applicants that are between 10-16 years old and that are 65 plus years old. Not having this data could not allow for the research to show how the future market, 10-16 year olds now, perceive chatbots also, how 65 plus year olds that also need support perceive chatbots is not known from this research and thus recommendations for them cannot be made.

9.2 Future Research

This research showed that the variables competence and empathy caused for a more likable chatbot, however these variables only signified a moderate causal relationship. This shows that there are other variables either related to humanness or unassociated variables that cause for a more likable chatbot. Future research regarding chatbots relating to its likability need to be focused around these unknown variables to further improve the workings of chatbots. Also further research using neuroscientific tools can help provide a better understanding of the internal responses people show when using a chatbot.

10. ACKNOWLEDGEMENTS

I would like to take this opportunity to show great respect and gratitude towards my supervisors Dr. C. Herrando and Dr. E. Constantinides for their aid during my research and the creation of this thesis. Furthermore, I would like to thank all the applicants that took part in the survey which was crucial for the realisation of this thesis. Thank you all.

11. REFERENCES

- Abdul-Kader, S. A. & Woods, J.C. (2015) Survey on Chatbot Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications*, 6 (7). ISSN 2156-5570
- Abu Shawar, B. & Atwell, E. (2007). Chatbots: Are they Really Useful? *LDV-Forum 2007 – Band 22 (1)* – 29-49
- Adiwardana, D., Luong, M. T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., ... & Le, Q. V. (2020). Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*.
- Airenti, G., Cruciani, M. & Plebe, A. (2019). The Cognitive Underpinnings of Anthropomorphism. *Frontiers in Psychology*.
- Algemene Zaken (2020, June 4). Dutch measures against coronavirus. Retrieved from <https://www.government.nl/topics/coronavirus-covid-19/tackling-new-coronavirus-in-the-netherlands>
- Banchs, R. E. (2017). On the construction of more human-like chatbots: Affect and emotion analysis of movie dialogue data. *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*.
- Bartneck, C., Kulic, D., Croft, E. & Zoghbi, S. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, 1, 71-81.
- Benbasat, I., Dimoka, A., Pavlou, P. A. & Qiu, L. (2019). The role of demographic similarity in people's decision to interact with online anthropomorphic recommendation agents: Evidence from a functional magnetic resonance imaging (fMRI) study *International Journal of Human-Computer Studies*
- Byrne, D. (1971) *The Attraction Paradigm*. New York: Academic Press.
- Carli, L. L., Ganley, R., & Pierce-Otay, A. (1991). Similarity and satisfaction in roommate relationships. *Personality and Social Psychology Bulletin*, 17, 419-426.
- Changchit, C. (2006). Consumer Perceptions of Online Shopping. *Issues in Information Systems Volume VII, No. 2, 2006*.
- Cialdini, R. B. (1993). *Influence: Science and practice* (3rd ed.). New York: Harper-Collins
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. doi:10.1037/0033-295x.114.4.864
- Fein, J., Morana, S. & Gnewuch, U. (2019). Measuring Service Encounter Satisfaction with Customer Service Chatbots using Sentiment Analysis. *Karlsruhe Institute of Technology, Institute of Information Systems and Marketing (IISM)*
- Go, E. & Shyam Sundar, S. (2019). Humanizing Chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* (2019).
- Hill, J., Randolph Ford, W. & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior*, 49, 245–250. doi:10.1016/j.chb.2015.02.026
- Hogg, M.A., Abrams, D. & Martin, G.N. (2010). Social cognition and attitudes. *Martin, G.N., Carlson, N.R., Buskist, W., (Ed.), Psychology (pp 646-677)*. Harlow: Pearson Education Limited.
- Hogg, M. A., Cooper-Shaw, L., & Holzworth, D. (1993). Group prototypicality and depersonalized

- attraction in small interactive groups. *Personality and Social Psychology Bulletin*, 19, 452-565.
- Jain, M., Kumar, P., Kota, R. & Patel, S. N. (2018). Evaluating and Informing the Design of Chatbots. *IBM Research, India; Realtor.com, Canada; Computer Science & Engineering, University of Washington, USA*
 - Jenkins, M. C., Churchill, R., Cox, S. & Smith, D. (2007) Analysis of User Interaction with Service Oriented Chatbot Systems. *Jacko J.A. (eds) Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments. HCI 2007. Lecture Notes in Computer Science*, vol 4552.
 - Kenrick, D. T., Neuberg, S. L., & Cialdini, R. B. (2002). Social psychology: Unraveling the mystery (2nd ed.). *Boston: Allyn & Bacon*
 - Kiesler, S., & Goetz, J. (2002). Mental models and cooperation with robotic assistants. In *Proceedings of CHI'02 on Human Factors in Computing Systems* (pp. 576–577). Minneapolis, MN: ACM Press
 - Morrissey K. & Kirakowski J. (2013) 'Realness' in Chatbots: Establishing Quantifiable Criteria. *Kurosu M. (eds) Human-Computer Interaction. Interaction Modalities and Techniques. HCI 2013. Lecture Notes in Computer Science*, vol 8007.
 - Moussawi, S., Koufaris, M., Benbunan-Fich, R. (2020) How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents *Institute of Applied Informatics at University of Leipzig 2020*
 - Pallant, J. (2010). SPSS survival manual: A step by step guide to data analysis using SPSS. *Open University Press/McGraw-Hill*
 - Pegasystems (2018). The chatbot revolution. *PEGA*
 - Pereira, J. & Díaz, O. (2018). A quality analysis of Facebook Messenger's most popular chatbots. *ACM SAC Conference, France, April 9-13, 2018 (SAC'18)*, 7 pages. DOI: 10.1145/3167132.3167362
 - Pijnenborg, G.H.M., Spikman, J.M., Jeronimus, B.F. & Aleman, A. (2012). Insight in schizophrenia: associations with empathy. *European Archives of Psychiatry and Clinical Neuroscience*. 263 (4): 299–307. doi:10.1007/s00406-012-0373-0
 - Roskos-Ewoldsen, D. R., Bichsel, J., & Hoffman, K. (2002). The Influence of Accessibility of Source Likability on Persuasion. *Journal of Experimental Social Psychology*, 38(2), 137–143. doi:10.1006/jesp.2001.1492
 - Shouse, E. (Dec. 2005) "Feeling, Emotion, Affect," *M/C Journal*, 8(6). Retrieved from <<http://journal.media-culture.org.au/0512/03-shouse.php>>.
 - Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European journal of operational research*, 177(3), 1333-1352.
 - Thomas, N.T. (2016). An E-business Chatbot using AIML and LSA. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*
 - Turing, A. M. (2009). Computing Machinery and Intelligence. *Parsing the Turing Test*, 23–65. doi:10.1007/978-1-4020-6710-5_3
 - Weinert, F. E. (2001). Concept of competence: A conceptual clarification. *D. S. Rychen & L. H. Salganik (Eds.), Defining and selecting key competencies* (p. 45–65). *Hogrefe & Huber Publishers*.
 - Weizenbaum, J. (1966). ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine. *Computational Linguistics, Association for Computing Machinery*, Vol. 9, No. 1, pp. 36-45. doi:10.1145/365153.365168
 - Xu, A., Liu, Z., Guo, Y., Sinha, V. & Akkiraju, R. (2017). A New Chatbot for Customer Service on Social Media. *IBM Research, Almaden; San Jose, USA*
 - Zue, V. W. & Glass, J. R. (2000). Conversational interfaces: advances and challenges. *Proceedings of the IEEE*, 88(8), 1166–1180. doi:10.1109/5.880078
 - Zumstein, D. & Hundertmark, S. (2017). Chatbots – An Interactive Technology for Personalized Communication, Transactions and Services. *IADIS International Journal on WWW/Internet*, Vol. 15, No. 1, pp. 96-109.

12. APPENDICES

12.1 Survey Design

Survey questionnaire (151 applicants):

1. Gender
 - a. Male.....48%
 - b. Female.....51,3%
 - c. other.....0,7%
2. Age
 - a. 10 – 15.....0%
 - b. 16 – 20.....12%
 - c. 21 – 25.....56,7%
 - d. 26 – 3523,3%
 - e. 36 – 507,3%
 - f. 51 – 640,7%
 - g. 65+
3. Education Level
 - a. Highschool Student.....15,3%
 - b. Bachelor Student.....52,7%
 - c. Master Student.....26%
 - d. Doctoral Student.....5,3%
 - e. Other.....0,7%
4. Employment Situation
 - a. Employed.....45%
 - b. Self-Employed.....12,8%
 - c. Unemployed.....26,2%
 - d. Retired.....0%
 - e. Other.....16,1%
5. Nationality
6. Do you have any experience with or knowledge of chatbots?
 - a. Yes.....80,1%
 - b. No.....19,9%
7. Chatbots always answer my questions.
8. The responses I get from chatbots make sense.
9. The chatbot does not contradict itself.
10. The responses from the chatbot are specific to my input.
11. I often get generic answers such as yes and ok.
12. I feel like my questions are not answered satisfactory.
13. The chatbot feels nice to talk to.
14. I get a feeling I can ask the chatbot anything.
15. I consider the chatbot a he/she instead of an it.
16. The chatbot shows emotions.
17. I can notice when the chatbot shows happiness in its response.
18. The chatbot notices my anger/happiness and replies accordingly.
19. I consider following advice given by chatbots.
20. I believe the information given by chatbots to be true.
21. I show interest in products recommended by chatbots.
22. I consider my experience with chatbots to be pleasant.
23. I would recommend chatbot usage to others.
24. I prefer chatbots over human chat.
25. E-Mail

Survey Items per variable:

Competence	Sensibleness	7. Chatbots always answer my questions. 8. The responses I get from chatbots make sense. 9. The chatbot does not contradict itself.
	Specificity	10. The responses from the chatbot are specific to my input. 11. I often get generic answers such as yes and ok. (reverse the likert scale values) 12. I feel like my questions are not answered satisfactory. (reverse the likert scale values)
Empathy	Affect	13. The chatbot feels nice to talk to. 14. I get a feeling I can ask the chatbot anything. 15. I consider the chat bot a he/she instead of an it .
	Emotion	16. The chatbot shows emotions. 17. I can notice when the chatbot shows happiness in its response. 18. The chatbot notices my anger/happiness and replies accordingly.
Likability		19. I consider my experience with chatbots to be pleasant. 20. I would recommend chatbot usage to others. 21. I prefer chatbots over human chat.

Consideration		22. I consider following advice given by chatbots. 23. I believe the information given by chatbots to be true. 24. I show interest in products recommended by chatbots.
---------------	--	---

12.2 Survey Data

Mean scores per variable:

Statistics

		Sensibleness	Specificity	Affect	Emotion	Consideration	Competence	Likability	Empathy
N	Valid	117	117	117	117	117	117	117	117
	Missing	34	34	34	34	34	34	34	34
Mean		4,0598	3,5926	2,9858	2,8262	4,0513	3,8262	3,3362	2,9060

Statistics

		"Chatbots always answer my questions."	"The responses I get from chatbots make sense."	"The chatbot does not contradict itself."	"The responses from the chatbot are specific to my input."	"I often get generic answers such as yes and ok."	"I feel like my questions are not answered satisfactory."	"The chatbot feels nice to talk to."	"I get a feeling I can ask the chatbot anything."	"I consider the chatbot a he/she instead of an it."	"The chatbot shows emotions."
N	Valid	117	117	117	117	117	117	117	117	117	117
	Missing	34	34	34	34	34	34	34	34	34	34
Mean		3,97	4,17	4,04	4,09	3,54	3,15	3,36	3,11	2,49	2,52

Statistics

		"I can notice when the chatbot shows happiness in its response."	"The chatbot notices my anger/happiness and replies accordingly."	"I consider following advice given by chatbots."	"I believe the information given by chatbots to be true"	"I show interest in products recommended by chatbots"	"I consider my experience with chatbots to be pleasant."	"I would recommend chatbot usage to others."	"I prefer chatbots over human chat."
N	Valid	117	117	117	117	117	117	117	117
	Missing	34	34	34	34	34	34	34	34
Mean		3,21	2,74	3,85	4,79	3,51	3,99	3,72	2,30

12.3 Correlation in data

Correlations

		Competence	Empathy	Likability	Consideration
Competence	Pearson Correlation	1	,329**	,628**	,593**
	Sig. (2-tailed)		,000	,000	,000
	N	117	117	117	117
Empathy	Pearson Correlation	,329**	1	,593**	,465**
	Sig. (2-tailed)	,000		,000	,000
	N	117	117	117	117
Likability	Pearson Correlation	,628**	,593**	1	,694**
	Sig. (2-tailed)	,000	,000		,000
	N	117	117	117	117
Consideration	Pearson Correlation	,593**	,465**	,694**	1
	Sig. (2-tailed)	,000	,000	,000	
	N	117	117	117	117

** . Correlation is significant at the 0.01 level (2-tailed).

12.4 Case Summaries

Case Summaries

Mean

Age	Likability
16 - 20	3,1818
21 - 25	3,4749
26 - 35	3,2564
36 - 50	2,4286
Total	3,3362

Case Summaries

Mean

Education Level	Likability
Bachelor Student	3,3990
Doctoral Student	3,2500
Highschool Student	3,2424
Master Student	3,2688
Other	3,0000
Total	3,3362

Case Summaries

Mean

Gender	Likability
Female	3,4943
Male	3,1667
Other	4,0000
Total	3,3362

Case Summaries

Mean

Employment Situation	Likability
	4,0000
Employed	3,1212
Other	3,5333
Self-Employed	3,2778
Unemployed	3,6092
Total	3,3362